4618–4625 *Nucleic Acids Research, 1998, Vol. 26, No. 20*

Mutational analysis of *Chlorella* virus DNA ligase: catalytic roles of domain I and motif VI

Verl Sriskanda and Stewart Shuman*

Molecular Biology Program, Sloan-Kettering Institute, 1275 York Avenue, New York, NY 10021, USA

Received July 1, 1998; Revised and Accepted August 26, 1998

ABSTRACT

A conserved catalytic core of the ATP-dependent DNA ligases is composed of an N-terminal domain (domain 1, containing nucleotidyl transferase motifs I, III, IIIa and IV) and a C-terminal domain (domain 2, containing motif VI) with an intervening cleft. Motif V links the two structural domains. Deletion analysis of the 298 amino acid *Chlorella* virus DNA ligase indicates that motif VI plays a critical role in the reaction of ligase with ATP to form ligase–adenylate, but is dispensable for the two subsequent steps in the ligation pathway; DNA–adenylate formation and strand closure. We find that formation of a phosphodiester at a pre-adenylated nick is subject to a rate limiting step that does not apply during the sealing of nicked DNA by ligase–adenylate. This step, presumably conformational, is accelerated or circumvented by deleting five amino acids of motif VI. The motif I lysine nucleophile (Lys27) is not required for strand closure by wild-type ligase, but this residue enhances the closure rate by a factor of 16 when motif VI is truncated. We find that a more extensively truncated ligase consisting of only N-terminal domain 1 and motif V is inert in ligase-–adenylate formation, but competent to catalyze strand closure at a pre-adenylated nick. These results suggest that different enzymic catalysts facilitate the three steps of the DNA ligase reaction.

INTRODUCTION

The DNA ligases of eukaryotes, eukaryotic viruses and bacterial viruses catalyze the joining of 5′-phosphate-terminated strands to 3′-hydroxyl-terminated strands via a common pathway involving three sequential nucleotidyl transfer reactions (1–4). In the first step, attack on the α-phosphate of ATP by ligase results in release of pyrophosphate and formation of a ligase–adenylate intermediate in which AMP is linked covalently to the ε-amino group of a lysine. The nucleotide is then transferred to the 5′-phosphate-terminated DNA strand to form a DNA–adenylate intermediate, A(5′)pp(5′)N. Attack by the 3′-OH strand on DNA–adenylate joins the two polynucleotides and liberates AMP. The first two steps of the ligase reaction are similar to the nucleotidyl transfer reactions performed by the eukaryotic mRNA capping enzymes (5,6). A lysine residue on the capping enzyme attacks the α-phosphate of GTP to form a covalent enzyme–GMP intermediate. The GMP is then transferred to the 5′-diphosphate terminus of RNA. The G(5′)ppp(5′)N capping reaction product resembles the DNA–adenylate intermediate in the ligase reaction.

The ATP-dependent ligases and GTP-dependent capping enzymes make up an enzyme superfamily defined by a set of six short motifs (I, III, IIIa, IV, V and VI) arrayed in the same order and with similar spacing (Fig. 1; 6–8). The sequence similarity between the polynucleotide ligases and the capping enzymes is limited to these segments, which suggested a common core structure in which the motifs are brought together at the enzyme active site. This prediction was borne out by the crystal structures of bacteriophage T7 DNA ligase with bound ATP and *Chlorella* virus capping enzyme with bound GTP (9,10). Both enzymes consist of a larger N-terminal domain (domain 1, containing motifs I, III, IIIa and IV) and a smaller C-terminal domain (domain 2, containing motif VI) with a cleft between them. Motif V serves as the bridging segment between the two domains.

The ATP binding site of T7 DNA ligase is made up of motifs I, III, IIIa, IV and V (9). The active site lysine residue to which AMP becomes covalently linked is located within conserved motif I (KxDGxR). In the T7 ligase crystal, the lysine nucleophile is positioned near the α-phosphate of ATP, while the motif I arginine forms a hydrogen bond with the 3′-OH of the ribose sugar of ATP. Other side chain interactions with ATP include: (i) a glutamate in motif III contacts the 2′-OH ribose sugar; (ii) a tyrosine in motif IIIa stacks on the adenine base; (iii) a lysine in motif V contacts the α-phosphate (9). Similar contacts between GTP and conserved functional groups in motifs I, III, IIIa and V are seen in the capping enzyme–GTP co-crystal (10).

We are examining the structure and mechanism of eukaryotic DNA ligases using vaccinia virus and *Chlorella* virus enzymes as models. We are especially interested in defining the moieties on the enzymes responsible for DNA nick recognition and catalysis of the three chemical steps of the ligation reaction. Both viral enzymes are capable of binding specifically to a nicked duplex DNA ligand containing a 3′-OH and a 5′-PO_4 at the nick (11–14). The 298 amino acid *Chlorella* virus PBCV-1 DNA ligase is the smallest eukaryotic DNA ligase known (13). PBCV-1 ligase contains no additional sequence downstream of motif VI, which is located at its extreme C-terminus (Fig. 1). It contains only 26 amino acids N-terminal of the active site (Lys27) within motif I.

*To whom correspondence should be addressed. Tel: +1 212 639 7145; Fax: +1 212 717 3623; Email: s-shuman@ski.mskcc.org

We therefore regard the *Chlorella* virus enzyme as the minimal catalytic unit of a eukaryotic DNA ligase.

Our aim is to generate a comprehensive structure–function map of PBCV-1 DNA ligase by targeted mutagenesis. Initial analysis of motif I ($_{27}$KxDGxR$_{32}$) revealed essential roles for the Lys, Asp and Arg side chains at different steps of the ligase reaction (14). Here, we focus on motif VI, which is conserved in eukaryotic ligases, an archaebacterial ligase and T7 ligase (Fig. 1). The distal half of motif VI contains a cluster of charged amino acids (consensus sequence RX$_{1–3}$D[K/R]) that is also present in motif VI of the mRNA capping enzymes. The crystal structure of *Chlorella* virus RNA capping enzyme shows that the Arg and Lys side chains of the RxDK element in motif VI interact with the β- and γ-phosphates of GTP (10). These contacts are specific to the 'closed' conformation of the protein–GTP complex, in which the interdomain cleft is narrowed by movement of the C-terminal domain toward the N-terminal domain (10). Mutational analysis of the yeast capping enzyme showed that the Arg, Asp and Lys side chains of the RxDK component of motif VI are essential for enzyme function *in vivo* (8). In the T7 DNA ligase crystal, the interdomain cleft is open and there are no contacts between the C-terminal domain (including motif VI) and ATP. However, the distal portion of motif VI ($_{350}$GTEDN$_{354}$; refer to Fig. 1) was poorly ordered in the T7 ligase crystal and was not included in the structural model (9).

We have examined the effects of deletion and missense mutations in motif VI on the activities of PBCV-1 DNA ligase. Our results indicate that the distal half of motif VI is important in the formation of ligase–adenylate, but dispensable for subsequent steps in the ligation pathway. We find that a truncated enzyme containing only domain I and motif V is able to catalyze strand closure at a pre-adenylated nick.

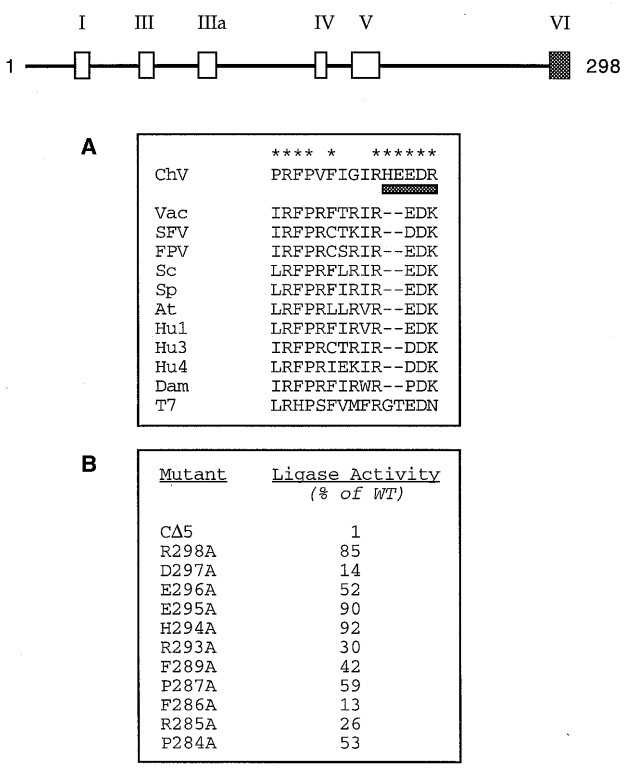

Figure 1. Mutational analysis of conserved motif VI. The 298 amino acid *Chlorella* virus DNA ligase polypeptide is depicted as a straight line, with the positions of conserved motifs I, III, IIIa, IV, V and VI denoted by boxes. (**A**) The sequence of motif VI of *Chlorella* virus ligase (ChV) is aligned with the corresponding sequences of the following ATP-dependent DNA ligases: vaccinia virus (Vac), Shope fibroma virus (SFV), fowlpox virus (FPV), *Saccharomyces cerevisiae* ligase I (Sc), *Schizosaccharomyces pombe* (Sp), *Arabidopsis thaliana* (At), human ligase I (Hu1), human ligase III (Hu3), human ligase IV, (Hu4), *Desulfurolobus ambivalens* (Dam) and bacteriophage T7 (T7). The five amino acid segment of ChV ligase deleted from mutant CΔ5 is underlined. Amino acid residues that were replaced individually by alanine are denoted by asterisks. (**B**) The specific activities of wild-type (WT) and mutant ligases were determined from the slopes of the titration curves in the linear range of enzyme dependence. Each specific activity value is the average of two titration experiments. The activity values for the mutant enzymes were normalized to the wild-type specific activity.

MATERIALS AND METHODS

Ligase mutants

Mutated versions of the PBCV-1 ligase gene encoding polypeptides truncated by five amino acids (CΔ5) or 98 amino acids (CΔ98) were PCR amplified from plasmid pET-ChV-ligase (13) using antisense oligonucleotide primers that introduced translation stop codons in lieu of the His293 and Thr201 codons. Missense mutations in the PBCV-1 ligase gene were programmed by synthetic oligonucleotides using the two-stage PCR-based overlap extension strategy (15). An *Nde*I–*Bam*HI restriction fragment of each PCR-amplified gene was inserted into pET16b so as to place the coding sequence in-frame with an N-terminal leader sequence encoding 10 tandem histidines. The presence of the desired mutation was confirmed in every case by sequencing the entire ligase insert; the occurrence of PCR-generated mutations outside the targeted region was thereby excluded. The pET-His-ligase plasmids were transformed into *Escherichia coli* BL21(DE3). IPTG-induced ligase expression and the purification of wild-type and mutant ligases from soluble bacterial extracts by Ni–NTA–agarose and phosphocellulose column chromatography were performed as described previously (14). The polypeptide composition of the column fractions was monitored by SDS–PAGE. The wild-type, CΔ5 and Ala-substituted ligases were step eluted from phosphocellulose between 0.2 and 0.4 M NaCl, whereas CΔ98 was eluted between 0.1 and 0.2 M NaCl. The protein concentrations of the enzyme preparations were determined using the BioRad dye reagent with bovine serum albumin as standard.

Ligation assay

The standard substrate used in ligase assays was a 36 bp DNA duplex containing a centrally placed nick (Fig. 2A). The 18mer constituting the 5′-phosphate-terminated strand d(ATTCCGATAGTGACTACA) was 5′-^{32}P-labeled and gel purified as described (7). The labeled 18mer was then annealed to a complementary 36mer in the presence of a 3′-OH 18mer strand d(CATATCCGTGTCGCCCTT) (11,12). Ligation reaction mixtures (20 µl) containing 50 mM Tris–HCl (pH 7.5), 5 mM DTT, 10 mM $MgCl_2$, 1 mM ATP, 5′-^{32}P-labeled nicked duplex substrate and enzyme were incubated at 22°C. Reactions were initiated by addition of enzyme and halted by addition of 1 µl 0.5 M EDTA and 5 µl formamide. The samples were heated at 95°C for 5 min and then analyzed by electrophoresis through a 17% polyacrylamide

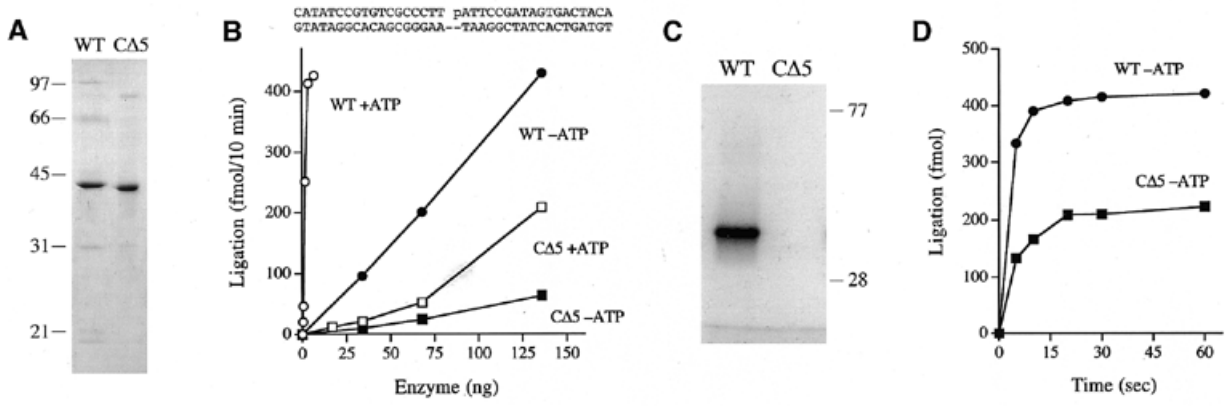

Figure 2. Characterization of CΔ5. (**A**) Ligase purification. Aliquots (0.8 µg) of the phophocellulose preparations of wild-type (WT) and CΔ5 ligase were analyzed by SDS–PAGE. The gel was stained with Coomassie blue dye. The positions and sizes (in kDa) of co-electrophoresed marker proteins are indicated on the left. (**B**) DNA ligation. The structure of the nicked duplex DNA substrate is shown. Ligation reaction mixtures containing 500 fmol nicked substrate, either 1 mM ATP (+ATP) or no added ATP (–ATP) and increasing amounts of wild-type or CΔ5 ligase were incubated for 10 min at 22°C. The yield of 36mer ligation product is plotted as a function of input enzyme. (**C**) Enzyme–adenylate formation. Reaction mixtures (20 µl) containing 50 mM Tris–HCl (pH 8.0), 5 mM DTT, 5 mM $MgCl_2$, 5 µM [α-^{32}P]ATP and 140 ng wild-type or CΔ5 ligase were incubated for 5 min at 37°C. Reactions were quenched by adding SDS to 1%. The reaction products were resolved by SDS–PAGE. An autoradiogram of the dried gel is shown. The positions and sizes (in kDa) of co-electrophoresed prestained marker proteins are indicated on the right. (**D**) Kinetic analysis of ATP-independent ligation. Reaction mixtures (20 µl) containing 500 fmol nicked duplex substrate and 8 pmol wild-type or CΔ5 ligase were incubated at 22°C. The extent of ligation (fmol 36mer product formed) is plotted as a function of reaction time.

gel containing 7 M urea. The extent of ligation [36mer/(18mer + 36mer)] was determined by scanning the gel using a Fujix BAS1000 PhosphorImager.

RESULTS

Effects of truncating motif VI on strand joining activity

A truncated version of PBCV-1 ligase, CΔ5, lacking the C-terminal pentapeptide HEEDR of motif VI (Fig. 1), was expressed in bacteria as an N-terminal His-tagged fusion protein. The wild-type ligase was expressed in parallel. Both proteins were purified from soluble lysates by Ni–agarose and phosphocellulose column chromatography. SDS–PAGE analysis showed that the CΔ5 protein migrated slightly faster than the wild-type enzyme, as expected (Fig. 2A). Mutational effects on ligation of a singly nicked duplex substrate were quantitated by enzyme titration in the presence of 1 mM ATP (Fig. 2B). The extent of ligation was proportional to input enzyme for the wild-type ligase; the reaction saturated with ~85% of the ^{32}P-labeled 18mer strand converted to 36mer in 10 min (Fig. 2B). This upper limit of ligation probably reflects incomplete annealing of all three component strands to form the nicked substrate. The specific activity of the CΔ5 mutant was 1% of the activity of the wild-type enzyme (Fig. 2B). Thus, the C-terminal pentapeptide of motif VI is important for enzyme function.

Strand joining by wild-type ligase could be detected in the absence of added ATP at stoichiometric levels of input enzyme (Fig. 2B, –ATP). ATP-independent ligation is attributable to ligase–adenylate in the enzyme preparation. The linear dependence of ATP-independent strand joining on enzyme indicated that 11% of the enzyme molecules had AMP bound at the active site. The activity of the CΔ5 mutant in the absence of ATP was much lower than that of the wild-type ligase (Fig. 2B); we estimate that 2% of the CΔ5 enzyme molecules were adenylated. Note that the specific activity of CΔ5 was stimulated only 2–3-fold by inclusion of 1 mM ATP, compared with >40-fold stimulation of the wild-type ligase by ATP (Fig. 2B).

To address the possibility that the ligase activity detected in the CΔ5 preparation might be caused by contamination with *E.coli* DNA ligase, we repeated the CΔ5 titration experiments in the presence and absence of 20 µM NAD. Control experiments using purified recombinant *E.coli* DNA ligase confirmed that 20 µM NAD was a saturating cofactor concentration (not shown). We found that the specific activity of CΔ5 in nick joining was not increased by the presence of NAD; this contrasts with the 100-fold stimulation by NAD of the specific activity of recombinant *E.coli* DNA ligase (data not shown). We surmise that the observed ligase activity is intrinsic to CΔ5.

Formation of the covalent ligase–adenylate intermediate *in vitro* was assayed by label transfer from [α-^{32}P]ATP to the enzyme (13). Incubation of wild-type ligase in the presence of 5 µM [α-^{32}P]ATP and a divalent cation resulted in formation of a ligase–AMP adduct that migrated as a single species during SDS–PAGE (Fig. 2C). The CΔ5 mutant was inactive in enzyme–AMP formation under these conditions (Fig. 2C). We presume that a defect in ligase adenylation (caused either by an increased K_m for ATP or a decreased rate of reaction chemistry, or both) contributes to the feeble stimulation by 1 mM ATP of the nick joining reaction of CΔ5.

Further insight into the basis for defective nick joining by CΔ5 emerged from an analysis of the kinetics of ligation in the absence of ATP. Under these circumstances, the reaction is limited to a single round of catalysis of steps 2 and 3 by pre-adenylated ligase. The reactions were performed in enzyme excess. As noted previously (14), single turnover ligation by wild-type PBCV-1 enzyme occurred rapidly; 79% of the end point was attained in 5 s (Fig. 2C). The salient finding was that single turnover ligation by CΔ5 was also rapid. The reaction was nearly complete in 20 s and attained 60% of the end point value in 5 s (Fig. 2C). The lower extent of ligation by CΔ5 versus wild-type in this experiment can be attributed to the lower percentage of pre-adenylated enzyme

in the preparation. These experiments show that deletion of five amino acids from motif VI inhibited ligase–adenylate formation, but had little effect on the catalysis of steps 2 and 3 by ligase–AMP.

DNA–adenylate formation on a gapped substrate

The second step of ligation is the transfer of AMP from ligase–adenylate to the 5′-phosphate terminus at the nick to form DNA–adenylate. This intermediate is not detected during ligation of nicked DNA by wild-type PBCV-1 ligase (13). However, DNA–adenylate accumulates to very high levels when wild-type ligase acts on a substrate containing a 1 nt gap between the reactive 3′-OH and 5′-phosphate DNA strands (14). In the experiment shown in Figure 3, reaction of the gapped substrate with an excess of wild-type DNA ligase in the presence of ATP resulted in conversion of the 5′-^{32}P-labeled 18mer strand into an adenylated species (AppDNA) that migrated ~1 nt slower than the input 18mer during polyacrylamide gel electrophoresis. Quantitation of the product distribution as a function of time showed that DNA–adenylate was more abundant than the 36mer ligation product during the linear phase of product accumulation (from 10 s to 1 min) and was not converted into ligated product with longer incubation (Fig. 4, WT+ATP). After 10 min, 45% of the labeled material was DNA–adenylate. In contrast, reaction of CΔ5 with the gapped substrate in the presence of ATP resulted in only a transient accumulation of DNA–adenylate at 30–60 s, followed by decay of AppDNA and a progressive accumulation of ligated 36mer (Fig. 3). After 30 min, 70% of the input gapped substrate was sealed and the ligated product was far more abundant than AppDNA (Fig. 4, CΔ5+ATP). It appeared that completion of step 3 by wild-type ligase under these reaction conditions was impeded compared with CΔ5.

An explanation for this effect emerged when we analyzed the reaction of the wild-type and CΔ5 enzymes with the 1 nt gapped substrate in the absence of ATP (Fig. 4). Under these conditions, DNA–adenylate formation by wild-type ligase peaked at 30–60 s (accounting for 21% of total labeled material) and decayed by 5 min, at which time most of the substrate was ligated (Fig. 4, WT–ATP). The ratio of AppDNA to ligated DNA at the wild-type ligase reaction end point, which was 1.7 in the presence of ATP, was reduced to 0.02 in the absence of ATP. Prior studies had shown that adenylated DNA ligase has a much lower binding affinity for a 1 nt gapped DNA than for a nicked DNA (13). Therefore, we propose the following scenario to account for the effects of ATP on DNA–adenylate accumulation: (i) wild-type ligase reacts with a nicked substrate to form nicked DNA–adenylate and immediately catalyzes strand closure without dissociating from the nicked DNA–adenylate intermediate; (ii) wild-type ligase reacts with the 1 nt gapped substrate to form gapped DNA–adenylate, which is prone to dissociate from the enzyme without progression through step 3; (iii) the free wild-type ligase that dissociates from the gapped DNA–adenylate is rapidly converted to ligase–adenylate in the presence of ATP and the adenylated enzyme is incapable of rebinding and reacting with DNA–adenylate, a situation resulting in the trapping of gapped DNA–adenylate; (iv) in the absence of ATP, the adenylate binding pocket of wild-type ligase remains free after step 2 and the enzyme can rebind to the gapped DNA–adenylate and catalyze strand closure across the gap. A prediction of this scheme is that a ligase mutant which is specifically defective in step 1 should not display ATP-dependent trapping of DNA–adenylate on the 1 nt gapped substrate. Such behavior is observed with CΔ5 (Fig. 4).

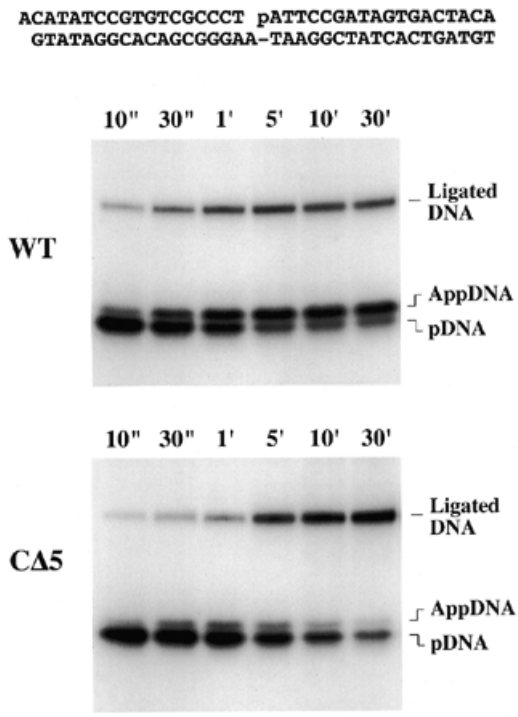

Figure 3. Ligase action at a 1 nt gap. The structure of the 1 nt gap substrate is shown. Reaction mixtures containing (per 20 µl) 50 mM Tris–HCl (pH 7.5), 5 mM DTT, 10 mM $MgCl_2$, 1 mM ATP, 500 fmol 1 nt gap DNA substrate and 6 pmol wild-type or CΔ5 ligase were incubated at 22°C. Aliquots (20 µl) were withdrawn at the times indicated above the lanes and quenched immediately with EDTA and formamide. The samples were analyzed by electrophoresis through a 20% polyacrylamide gel containing 7 M urea. An autoradiogram of the gel is shown. The positions of the input 5′ monophosphate 18mer strand (pDNA), the adenylated DNA strand (AppDNA) and the 36mer ligation product are shown on the right.

Phosphodiester bond formation at a pre-adenylated nick

Step 3 of the ligation reaction was assayed by the ability of wild-type enzyme and CΔ5 to seal a pre-adenylated nicked duplex DNA. The adenylated strand used to form this substrate was synthesized by ligase-mediated AMP transfer to the 5′-^{32}P-labeled strand of a DNA molecule containing a 1 nt gap. The radiolabeled AppDNA strand was purified by gel electrophoresis and annealed to an unlabeled 36mer template oligonucleotide and a 3′-OH 18mer oligonucleotide to form the structure shown in Figure 5A. This substrate was reacted for 10 min with PBCV-1 ligase in the presence of magnesium, without added ATP. Only the unadenylated form of ligase should be reactive in this assay. Control reactions containing the standard nicked DNA duplex were performed in parallel; in these reactions, only pre-adenylated ligase will catalyze strand closure. As an added control for these experiments, we tested the activity of the active site lysine mutant K27A in sealing nicked DNA–adenylate and nicked DNA substrates. The K27A enzyme cannot catalyze step 1 and therefore should contain no pre-adenylated ligase in the preparation (14).

The wild-type ligase joined the ^{32}P-labeled AppDNA oligonucleotide to the unlabeled 3′-OH 18mer to generate a

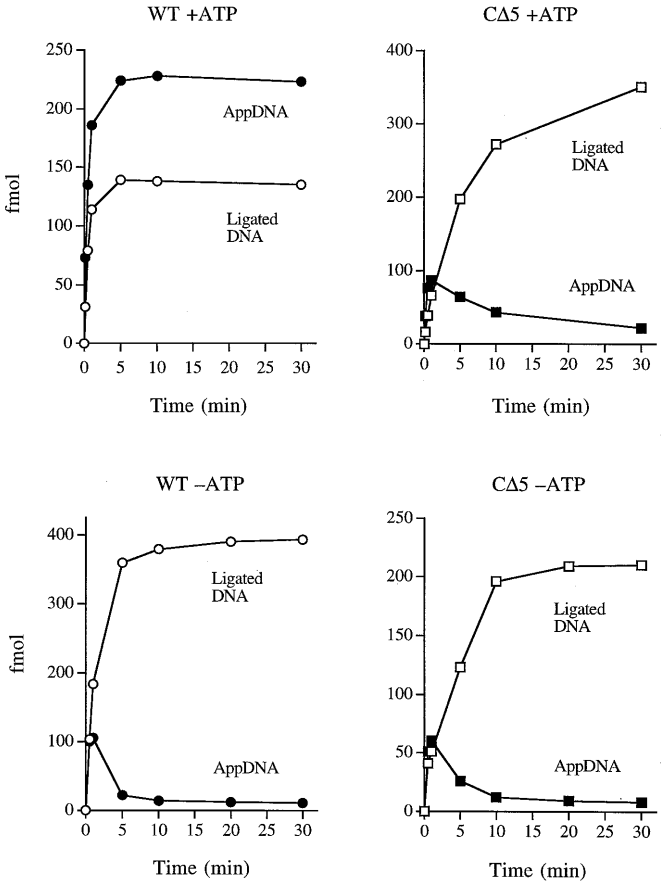

Figure 4. Effect of ATP on formation of gapped DNA–adenylate. Wild-type and CΔ5 ligase were reacted with the 1 nt gapped substrate as described in the legend to Figure 3. Reaction mixtures either contained 1 mM ATP (+ATP; top) or no added ATP (–ATP; bottom) as indicated. Reaction products were resolved by gel electrophoresis and the distribution of ^{32}P label as ligated DNA (36mer product), DNA–adenylate (AppDNA) and residual 18mer substrate was determined by scanning the gel with a PhosphorImager. The levels of ligated product and DNA–adenylate are plotted as a function of reaction time.

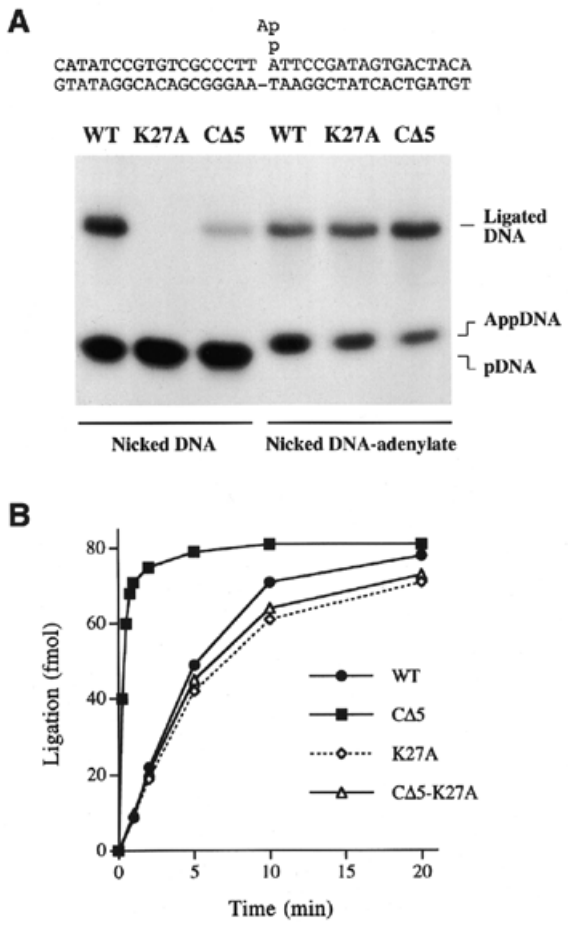

Figure 5. Strand joining at a pre-adenylated nick. (**A**) The structure of the nicked DNA–adenylate substrate is shown. The 5′-adenylated ^{32}P-labeled 18mer strand was synthesized and gel purified as described (6,7). Strand joining reaction mixtures (20 µl) containing 50 mM Tris–HCl (pH 7.5), 5 mM DTT, 10 mM $MgCl_2$, 100 fmol nicked DNA substrate (shown in Fig. 2) or nicked DNA–adenylate substrate and 200 fmol wild-type, K27A or CΔ5 ligase were incubated for 10 min at 22°C. The reaction products were resolved by denaturing polyacrylamide gel electrophoresis. An autoradiogram of the gel is shown. (**B**) Kinetics of sealing at a pre-adenylated nick. Reaction mixtures containing (per 20 µl) 100 fmol nicked DNA–adenylate substrate and 400 fmol wild-type, K27A, CΔ5 or CΔ5-K27A ligase were incubated at 22°C. Aliquots (20 µl) were withdrawn at the times indicated and quenched immediately. The extent of ligation (fmol 36mer product formed) is plotted as a function of reaction time.

^{32}P-labeled 36mer ligation product (Fig. 5A). CΔ5, which was less active than wild-type enzyme on a nicked duplex (because of the lower level of ligase–AMP in the preparation), was equally if not more active that the wild-type in sealing the pre-adenylated nick (Fig. 5A). K27A was unable to seal a nicked duplex (as expected), but did seal nicked DNA–adenylate. We did not detect conversion of the ^{32}P-labeled AppDNA oligonucleotide to 5′-^{32}P-monophosphate 18mer during the reaction of wild-type or mutant ligases with nicked DNA–adenylate.

The results of a kinetic analysis of phosphodiester bond formation at a pre-adenylated nick under conditions of ligase excess (4:1 molar ratio of ligase to DNA) were surprising and instructive (Fig. 5B). First, although 80% of the pre-adenylated nick was sealed by wild-type ligase (similar to the yield of ligated product on standard nicked DNA), the rate of reaction was slow compared with the rate of single turnover ligation of nicked DNA in enzyme excess. Joining of nicked DNA–adenylate was linear for 5 min and nearly complete at 10 min (Fig. 5B), whereas ATP-independent joining of nicked DNA was complete within 10 s (Fig. 2C). The rate of joining of nicked DNA–adenylate did not increase when the concentration of input ligase was increased to attain a 10:1 molar excess over the substrate (not shown). These results were puzzling, insofar as an enzyme ought not to be able to catalyze two sequential steps in a reaction pathway any faster than it catalyzes one of the steps in isolation.

In order to account for these findings, we posit that the reaction of wild-type ligase with the exogenous nicked DNA–adenylate substrate is subject to a rate limiting step that is slower than the rate limiting step of single turnover nick ligation. The novel rate limiting step in sealing exogenous nicked DNA–adenylate may entail a conformational change in the ligase or the DNA substrate. Our reasoning is as follows. During the joining reaction on a nicked DNA substrate, the DNA–adenylate intermediate is formed from ligase–adenylate bound at the nick and strand

closure is likely to proceed directly, i.e. the extrahelical nucleotide of nicked DNA–adenylate is already positioned correctly in the adenylate binding pocket on the enzyme. In contrast, the conformation of the 5′-adenylate moiety of the exogenous nicked DNA–adenylate substrate may be far from optimal. For example, Subramanya *et al*. (9) have suggested that the adenosine may intercalate into the nick unless held outside the helix by the ligase. Sealing of nicked DNA–adenylate might then be slowed by the need to flip the adenosine out of an unfavorable conformation. It is not clear if the enzyme itself facilitates this step.

The second remarkable finding was that ligation of nicked DNA–adenylate by CΔ5 was 16-fold faster than the wild-type rate (Fig. 5B). Indeed, the CΔ5 reaction with nicked DNA–adenylate attained half the end point value in 15 s (Fig. 5B), which was not much different from the behavior of CΔ5 in single turnover reaction with nicked DNA (Fig. 1C). This result implies that structural moieties on the wild-type enzyme actually impede the rate of reaction with exogenous nicked DNA–adenylate. Deleting the C-terminal five amino acids of motif VI appears to reverse the impediment and elicit a gain-of-function in this assay.

The active site lysine mutant K27A ligated the nicked DNA–adenylate substrate at the same rate as the wild-type enzyme (Fig. 5B). This result argues that loss of the lysine nucleophile does not affect the step that is rate limiting in the isolated step 3 reaction. However, it is difficult to draw a firm conclusion about the potential role of the active site lysine in step 3 chemistry, given that the reaction being measured is likely limited by pre-chemical steps. To approach this issue, we introduced the K27A mutation into the truncated CΔ5 polypeptide to yield a CΔ5-K27A double mutant. The CΔ5-K27A protein ligated the nicked DNA–adenylate substrate at the same rate as the wild-type and K27A proteins, but was slower by a factor of 16 than the CΔ5 enzyme (Fig. 5B). Thus, the active site lysine, though not strictly essential for step 3, clearly does contribute to the enzyme-mediated rate enhancement of phosphodiester bond formation.

Alanine scanning mutagenesis

Single alanine substitutions were introduced at 11 positions within and flanking motif VI of PBCV-1 DNA ligase, The residues mutated are denoted by asterisks in Figure 1. The Ala-substituted proteins were expressed in bacteria as His-tagged derivatives and purified from soluble bacterial lysates by Ni–agarose and phosphocellulose column chromatography. SDS–PAGE analysis showed that the purity of the the alanine mutants was equivalent to that of the wild-type enzyme depicted in Figure 2A (data not shown). Each protein was assayed by its ability to ligate nicked DNA in the presence of ATP. The specific activities of the mutants are tabulated in Figure 1. Whereas deleting the C-terminal five amino acids $_{294}$HEEDR$_{298}$ reduced specific activity by two orders of magnitude, elimination of the individual side chains had more modest effects, the most significant being the 7-fold decrement in specific activity elicited by the D297A mutation. The E296A mutation reduced activity by half, whereas the H284A, E295A and R298A mutations were without effect. Elimination of the conserved Arg285, Phe286 and Arg293 side chains reduced activity to 26, 13 and 30% of wild-type, respectively (Fig. 1).

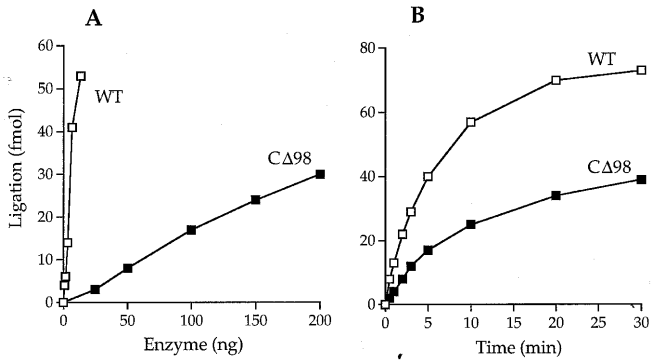

Figure 6. Characterization of CΔ98. (**A**) Strand joining at a pre-adenylated nick. Reaction mixtures (20 µl) containing 50 mM Tris–HCl (pH 7.5), 5 mM DTT, 10 mM $MgCl_2$, 100 fmol nicked DNA–adenylate and wild-type or CΔ5 ligase as indicated were incubated for 30 min at 22°C. The extent of ligation (fmol 36mer product formed) is plotted as a function of input ligase. (**B**) Kinetics. Reaction mixtures containing (per 20 µl) 100 fmol nicked DNA–adenylate and either 34 ng (1 pmol) wild-type ligase or 250 ng (10 pmol) CΔ5 ligase were incubated at 22°C. Aliquots (20 µl) were withdrawn at the times indicated and quenched immediately. The extent of ligation is plotted as a function of reaction time.

Effects of deleting domain 2

Structural domain 2 of T7 DNA ligase extends from the distal margin of motif V to the C-terminus. To assess the role of domain 2 in ligase function, we engineered a more extensively truncated mutant, CΔ98, which includes motifs I, III, IIIa, IV and V. The CΔ98 polypeptide (amino acids 1–200) terminates four residues downstream of motif V. CΔ98 was expressed in bacteria as an N-terminal His-tagged fusion protein and purified by Ni–NTA–agarose and phosphocellulose chromatography. CΔ98 was incapable of sealing a nicked duplex substrate in the presence of 1 mM ATP or in the absence of ATP, even when the truncated enzyme was added in 20-fold molar excess over the nicked substrate (not shown). CΔ98 also formed no detectable ligase–AMP adduct when reacted with [α-^{32}P]ATP *in vitro* (not shown). Thus, CΔ98 was apparently incapable of catalyzing step 1, even during its exposure to ATP during expression *in vivo*. CΔ98 did catalyze phosphodiester bond formation on the pre-adenylated nicked substrate, albeit with rather low efficiency on a per enzyme basis, as shown by the protein titration experiment shown in Figure 6A. The specific activity of CΔ98 in closing nicked DNA–adenylate was ~2% of the wild-type activity. A kinetic analysis of the strand closure reaction at the highest concentration of input CΔ98 showed that the rate was comparable with that of wild-type enzyme (Fig. 6B). These results suggest that CΔ98 is catalytically competent for step 3, but its specific activity is limited by virtue of severely diminished affinity for the nicked DNA–adenylate substrate. This would be in keeping with the suggestion of Doherty *et al*. (18) that domain 2 of T7 ligase comprises part of the DNA binding surface.

In order to ensure that the observed strand closure activity at a pre-adenylated nick was intrinsic to CΔ98 (as opposed to trace contamination with the 671 amino acid bacterial ligase), we sedimented the phosphocellulose preparation of CΔ98 through a glycerol gradient. Two peaks of strand joining activity were noted (Fig. 7, bottom). The activity profile coincided with the

sedimentation profile of the CΔ98 polypeptide (Fig. 7, top). The apparent size of CΔ98 gauged by SDS–PAGE (31 kDa) was greater than its calculated size of 25 kDa. This is in keeping with the anomalously slow electrophoretic mobility noted for the full-length PBCV-1 ligase. The 'light' component of CΔ98 sedimented at 2.6 S relative to marker proteins (catalase, 11.2 S, 248 kDa; BSA, 4.4 S, 66 kDa; cytochrome c, 1.9 S, 13 kDa) that were sedimented in a parallel gradient. We surmise that the 2.6 S component is a CΔ98 monomer. (*Escherichia coli* DNA ligase sediments as a 3.9 S monomer; 2.) The 'heavy' component sedimented at 9.0 S (~190 kDa), which suggests an octamer of the CΔ98 polypeptide, assuming that the oligomer is globular in shape. The bimodal sedimentation profile of CΔ98 contrasts with that of the full-length PBCV-1 ligase, which sediments as a single monomeric peak (13; data not shown). We verified that the 31 kDa polypeptides detected in glycerol gradient fractions 11 and 23 were indeed the His-tagged PBCV-1 protein by subjecting both species to automated Edman sequencing after transferring the polypeptides from an SDS gel to a polyvinylidene difluoride membrane. The experimentally determined sequences (xHHHHH) corresponded to residues 2–7 of the N-terminal leader peptide. We conclude from these experiments that CΔ98 is responsible for the observed strand closure activity.

DISCUSSION

The results presented above provide insights into catalysis by ATP-dependent DNA ligases, to wit: (i) deletion analysis indicates an important role for motif VI in ligase–adenylate formation; (ii) formation of a phosphodiester at a pre-adenylated nick is subject to a rate limiting step that is accelerated or circumvented by deleting the distal half of motif VI; (iii) the motif I lysine (Lys27), though not absolutely required for step 3, enhances the rate of strand closure by a factor of 16; (iv) a polypeptide consisting of domain 1 and motif V is sufficient to catalyze step 3 of the ligation pathway; (v) reaction of wild-type ligase with a gapped substrate in the presence of ATP results in formation of high levels of gapped DNA–adenylate.

Catalytic role of motif VI

The effect of removing the C-terminal five amino acids of motif VI was to reduce ligase specific activity by a factor of 10^{-2}. The level of endogenous ligase–adenylate in the CΔ5 enzyme preparation was 5-fold lower than in wild-type ligase and the purified CΔ5 enzyme was defective in ligase–AMP formation *in vitro*. Yet, the C-terminal deletion had no significant effect on the ability of pre-formed ligase–adenylate to seal a nick. This argues that the deleted five amino acid segment of motif VI is uniquely required for step 1 of the ligase reaction. We were unable, based on the effects of single alanine substitutions in the deleted peptide, to attribute the deletional effects to a single catalytic side chain.

A clue to how motif VI might function during the ligase adenylation step comes from crystallographic analysis of the capping enzyme–GTP complex. The capping enzyme asymmetric unit contains two GTP-bound enzymes with different conformations, open and closed, that differ by rigid body movement of domain 2 toward domain 1 (10). Conformational closure brings motif VI of capping enzyme into contact with the β- and γ-phosphates of GTP and reorients the phosphates so that the pyrophosphate leaving group is positioned apically to the lysine nucleophile. This facilitates in-line attack on the α-phosphorus of GTP. In the open conformation of the enzyme–GTP complex, the

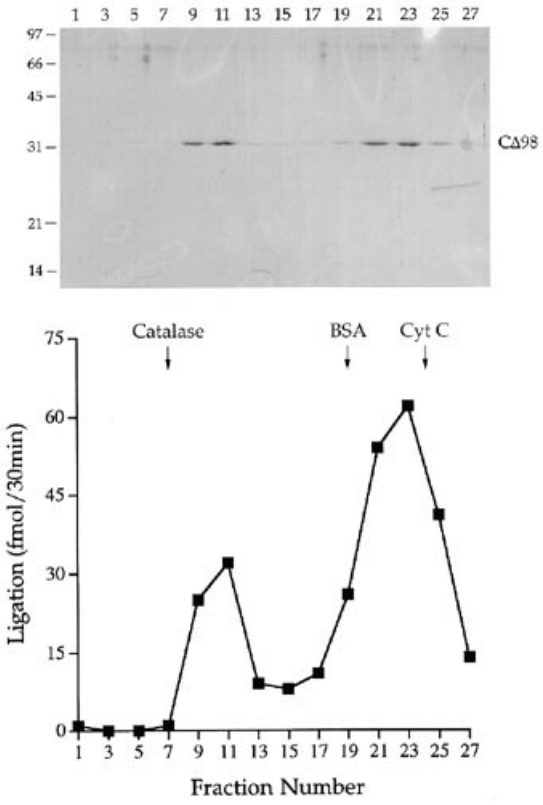

Figure 7. Sedimentation analysis of CΔ98. An aliquot (0.2 ml, 40 μg protein) of the phosphocellulose CΔ98 enzyme fraction was applied to a 4.8 ml 15–30% glycerol gradient containing 50 mM Tris–HCl (pH 8.0), 0.2 M NaCl, 1 mM EDTA, 2 mM DTT and 0.1% Triton X-100. The gradient was centrifuged at 4°C for 14 h at 50 000 r.p.m. in a SW50 rotor. Fractions were collected from the bottom of the tube. Aliquots (20 μl) of odd numbered gradient fractions were analyzed by SDS–PAGE. A Coomassie blue stained gel is shown in the top panel. The positions and sizes (in kDa) of co-electrophoresed size markers are indicated on the left. The position of CΔ98 is denoted on the right. The gradient fractions were assayed for strand joining activity at a pre-adenylated nick as described in Figure 6. The reaction mixtures contained 200 fmol nicked DNA–adenylate and 2 μl specified gradient fraction. The activity profile is shown in the bottom panel. The vertical arrows indicate the sedimentation peaks of marker proteins (catalase, BSA and cytochrome c) that were centrifuged in a parallel gradient.

β- and γ-phosphates are nearly orthogonal to the lysine, a configuration that is most unfavorable for catalysis (10). In the T7 ligase crystal with bound ATP, the β- and γ-phosphates of ATP are positioned orthogonal to the active site lysine (9). The C-terminal peptide of T7 ligase that aligns with the deleted five amino acid segment of the PBCV-1 CΔ5 mutant is disordered in the T7 ligase crystal (9). By analogy with the capping enzyme, we posit a conformation of DNA ligase in which motif VI interacts with the β- and γ-phosphates of ATP to orient them for in-line attack by Lys27. This model would explain the finding that deletion of PBCV-1 motif VI is deleterious to ligase–adenylate formation.

We envision that breakage of the α–β phosphoanhydride linkage of ATP during step 1 allows the ligase to revert to an open conformation that can accommodate nicked duplex DNA within the interdomain groove. Binding of nicked DNA depends on occupancy of the adenylate binding pocket on the enzyme and the presence of a 5′-phosphate at the nick (12–14). Catalysis of DNA–adenylate formation and strand closure would necessarily

be performed by enzymic functional groups located near the adenylate and the 5′-phosphate and 3′-OH moieties at the nick. We suspect that motif VI, if it contacts the DNA at all, would be positioned at a distance from the active site in the ligase–DNA complex. This is consistent with our observation that the CΔ5 deletion has little effect on the rate of single turnover ligation by preformed ligase–AMP. A corollary prediction is that only an open conformation of ligase–AMP can bind productively to nicked DNA and an open complex of unadenylated ligase bind to nicked DNA–adenylate. The CΔ98 protein, which retains the ability to catalyze step 3 on exogenous nicked DNA–adenylate, can be viewed as the extreme case of an open conformation, i.e. it effectively lacks all of domain 2 downstream of motif V.

Ligase action at a pre-adenylated nick

The rate of ligation of a pre-adenylated nick by CΔ5 was ~16-fold faster than the wild-type rate. We surmise that loss of the distal five amino acids of motif VI overcomes a rate limiting step that is unique to the reaction of ligase with an exogenous nicked DNA–adenylate substrate. We discussed above the possible requirement for a conformational change in the DNA substrate. It is also conceivable that conformational dynamics of the ligase influence the step 3 reaction. For example, the presence of motif VI in wild-type ligase may promote adoption of a conformation that is more favorable for ATP binding versus binding to nicked DNA–adenylate, in which case the CΔ5 deletion might alter the conformational equilibrium toward a state that binds to the adenylated DNA. Defining these conformational states will necessarily hinge on crystallizing the PBCV-1 ligase and bound intermediates at each step along the reaction pathway.

A prior study by Yang and Chan (16) documented the ability of human DNA ligases I and II to catalyze the sealing of an exogenous nicked DNA–adenylate substrate. Sealing of nicked DNA–adenylate by human ligases I required a divalent cation (16); this is also the case for PBCV-1 ligase and vaccinia ligase. The kinetics of sealing of nicked DNA–adenylate by human DNA ligase I were similar to results presented here for PBCV-1 ligase, i.e. the reaction was complete after 10 min (16). A key distinction is that reaction of human DNA ligases I and II with nicked DNA–adenylate resulted not only in strand closure (step 3), but also in substantial conversion of the 5′-adenylated strand (AppDNA) to a 5′-monophosphate (reversal of step 2). Indeed, DNA ligase II generated nearly equal amounts of 5′-monophosphate strand and ligated strand products when reacted with nicked DNA–adenylate (16). In contrast, the reaction of PBCV-1 DNA ligase with nicked DNA–adenylate yielded only ligated 36mer and there was no detectable formation of a deadenylated 5′-phosphate 18mer strand by reversal of step 2. Vaccinia DNA ligase resembles the *Chlorella* virus enzyme in this respect (J.Sekiguchi and S.Shuman, unpublished results). We surmise that members of the eukaryotic DNA ligase family differ significantly with respect to the relative rates of the reverse step 2 and step 3 reactions.

Modrich and Lehman (17) conducted a thorough steady-state kinetic analysis of the composite ligation reaction and the three partial reactions catalyzed by *E.coli* DNA ligase using homopolymeric substrates consisting of reactive poly(dT) strands annealed to poly(dA). They found that the turnover number for phosphodiester bond formation on a pre-adenylated AppT$(dT)_n$ substrate was lower than the turnover number on a nicked substrate under optimal conditions. Thus, step 3 on an exogenous DNA–adenylate is anomalously slower than the multistep reaction, whether catalyzed by the NAD-dependent bacterial ligase or the ATP-dependent *Chlorella* virus ligase. As did we, Modrich and Lehman invoked a slow conformational step for joining exogenous DNA–adenylate that need not occur when the reaction proceeds through the normal series of partial reactions (17). They also considered that reversal of step 2 during reaction of *E.coli* ligase with added DNA–adenylate might account in part for the slow step 3 rate. Reversal of step 2 appears not to be a factor in the step 3 joining reaction of PBCV-1 ligase.

Role of the motif I lysine in strand closure

The present results sound a cautionary note in the interpretation of mutational effects on joining nicked DNA–adenylate, to wit, that ligase mutations that slow the chemical step of phosphodiester formation by an order of magnitude, but not to the point that chemistry becomes rate limiting in the reaction with exogenous nicked DNA–adenylate, may be perceived by the investigator as having no impact on the reaction. This was the case presently for K27A. Our earlier conclusion that the active site nucleophile is not strictly essential for step 3 of the catalytic pathway still holds, but it is now apparent from analysis of the CΔ5-K27A double mutant that the lysine side chain contributes a 16-fold rate enhancement. The role of Lys27 in step 3 most likely entails interaction of the basic side chain with the α-phosphate of the adenylate of the AppDNA strand. It is possible that the lysine serves as a general acid to promote expulsion of the AMP leaving group during attack of the 3′-OH on DNA–adenylate. Different members of the polynucleotide ligase family may be more or less dependent on their respective motif I lysines to carry out step 3. Vaccinia and *Chlorella* virus DNA ligases still catalyze strand closure on nicked DNA adenylate when their motif I lysines are mutated to alanine (12,14). On the other hand, replacement of the active site lysine of T4 RNA ligase by asparagine inactivates the joining step, as assayed by the reaction of AAG with AppGp to form AAGG (19).

REFERENCES

1 Lehman,I.R. (1974) *Science*, **186**, 790–797.
2 Engler,M.J. and Richardson,C.C. (1982) *Enzymes*, **15**, 3–29.
3 Lindahl,T. and Barnes,D.E. (1992) *Annu. Rev. Biochem.*, **61**, 251–281.
4 Tomkinson,A.E. and Levin,D.S. (1997) *BioEssays*, **19**, 893–901.
5 Shuman,S. and Hurwitz,J. (1981) *Proc. Natl Acad. Sci. USA*, **78**, 187–191.
6 Shuman,S. and Schwer,B. (1995) *Mol. Microbiol.*, **17**, 405–410.
7 Shuman,S. and Ru,X. (1995) *Virology*, **211**, 73–83.
8 Wang,S.P., Deng,L., Ho,C.K. and Shuman,S. (1997) *Proc. Natl Acad. Sci. USA*, **94**, 9573–9578.
9 Subramanya,H.S., Doherty,A.J., Ashford,S.R. and Wigley,D.B. (1996) *Cell*, **85**, 607–615.
10 Hakansson,K., Doherty,A.J., Shuman,S. and Wigley,D.B. (1997) *Cell*, **89**, 545–553.
11 Shuman,S. (1995) *Biochemistry*, **34**, 16138–16147.
12 Sekiguchi,J. and Shuman,S. (1997) *J. Virol.*, **71**, 9679–9684.
13 Ho,C.K., Van Etten,J.L. and Shuman,S. (1997) *J. Virol.*, **71**, 1931–1937.
14 Sriskanda,V. and Shuman,S. (1998) *Nucleic Acids Res.*, **26**, 525–531.
15 Ho,S.N., Hunt,H.D., Horton,R.M., Pullen,J.K. and Pease,L.R. (1989) *Gene*, **77**, 51–59.
16 Yang,S. and Chan,J.Y.H. (1992) *J. Biol. Chem.*, **267**, 8117–8122.
17 Modrich,P. and Lehman,I.R. (1973) *J. Biol. Chem.*, **248**, 7502–7511.
18 Doherty,A.J., Ashford,S.R. and Wigley,D.B. (1996) *Nucleic Acids Res.*, **24**, 2282–2287.
19 Heaphy,S., Singh,M. and Gait,M.J. (1987) *Biochemistry*, **26**, 1688–1696.